



1/48

ATGGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGGCGCGGGAG  
TGCTGCCTGCCCACGGCACCCAGCACGGCATCCGGCTGCCCCTGCG  
CAGCGGCCTGGGGGGCGCCCCCTGGGGCTGCGGCTGCCCCGGGA  
GACCGACGAAGAGCCCCGAGGAGCCCCGGCCGGAGGGGCGAGCTTTGT  
GGAGATGGTGGACAACCTGAGGGGGCAAGTCGGGGCAGGGCTACTAC  
GTGGAGATGACCGTGGGCAGCCCCCGCAGACGCTCAACATCCTGG  
TGGATACAGGCAGCAGTAACTTTGCAGTGGGTGCTGCCCCCACCC  
CTTCCTGCATCGCTACTACCAGAGGCAGCTGTCCAGCACATAACGGG  
ACCTCCGGAAGGGTGTGTATGTGCCCTACACCCAGGGCAAGTGGGA  
AGGGGAGCTGGGCACCGACCTGGTAAGCATCCCCCATGGCCCCAAC  
GTCATGTGCGTGCCAACATTGCTGCCATCACTGAATCAGACAAGTT  
CTTCATCAACGGCTCCAACCTGGGAAGGCATCCTGGGGCTGGCCTATG  
CTGAGATTGCCAGGCCTGACGACTCCCTGGAGCCTTTCTTTGACTCT  
CTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAGCTTTG  
TGGTGCTGGCTTCCCCCTCAACCAGTCTGAAGTGCTGGCCTCTGTGCG  
GAGGGAGCATGATCATTGGAGGTATCGACCACTCGCTGTACACAGGC  
AGTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTATGAGGTGAT  
CATTGTGCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCA  
AGGAGTACAACTATGACAAGAGCATTGTGGACAGTGGCACCACCAAC  
CTTCGTTTGCCCAAGAAAGTGTTTGAAGCTGCAGTCAAATCCATCAAG  
GCAGCCTCCTCCACGGAGAAGTTCCCTGATGGTTTCTGGCTAGGAGA  
GCAGCTGGTGTGCTGGCAAGCAGGCACCACCCCTTGGAACATTTTCC  
CAGTCATCTCACTCTACCTAATGGGTGAGGTTACCAACCAGTCCTTCC  
GCATCACCATCCTTCCGCAGCAATACCTGCGGCCAGTGGAAGATGTG  
GCCACGTCCCAAGACGACTGTTACAAGTTTGCCATCTCACAGTCATC  
CACGGGCACTGTTATGGGAGCTGTTATCATGGAGGGCTTCTACGTTG  
TCTTTGATCGGGCCCCGAAAACGAATTGGCTTTGCTGTCAGCGCTTGC  
CATGTGCACGATGAGTTCAGGACGGCAGCGGTGGAAGGCCCTTTTG  
TCACCTTGGACATGGAAGACTGTGGCTACAACATTCCACAGACAGAT  
GAGTCAACCCTCATGACCATAGCCTATGTCATGGCTGCCATCTGCGC  
CCTCTTCATGCTGCCACTCTGCCTCATGGTGTGTCAGTGGCGCTGCC  
TCCGCTGCCTGCGCCAGCAGCATGATGACTTTGCTGATGACATCTCC  
CTGCTGAAG

FIG. 1A

**2/48**

CCATGCCGGCCCCCTCACAGCCCCGCGGGAGCCCCGAGCCCCGCTGCCCCAGG  
CTGGCCGCGCGSGTGCCGATGTAGCGGGCTCCGGATCCCAGCCTCTCCCCT  
GCTCCCGTGCTCTGCGGATCTCCCCTGACCGCTCTCCACAGCCCCGGACCCG  
GGGGCTGGCCCAGGGCCCTGCAGGCCCTGGCGTCCTGATGCCCCCAAGCT  
CCCTCTCCTGAGAAGCCACCAGCACCCAGACTTGGGGGCAGGCGCCA  
GGGACGGACGTGGGCCAGTGCGAGCCCAGAGGGCCCCGAAGGCCGGGGCC  
CACCATGGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGGCGCGGGAG  
TGCTGCCTGCCCACGGCACCCAGCACGGCATCCGGCTGCCCCCTGCGCAGC  
GGCCTGGGGGGGCGCCCCCTGGGGCTGCGGCTGCCCCGGGAGACCGACG  
AAGAGCCCCGAGGAGCCCCGCGCGGAGGGGCAGCTTTGTGGAGATGGTGGAC  
AACCTGAGGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACCGTGGG  
CAGCCCCCGCAGACGCTCAACATCCTGGTGGATACAGGCAGCAGTAACTT  
TGCAGTGGGTGCTGCCCCCCACCCCTTCTGCATCGCTACTACCAGAGGCA  
GCTGTCCAGCACATACCGGGACCTCCGGAAGGGTGTGTATGTGCCCTACAC  
CCAGGGCAAGTGGGAAGGGGAGCTGGGCACCGACCTGGTAAGCATCCCC  
ATGGCCCCAACGTCACCTGTGCGTGCCAACATTGCTGCCATCACTGAATCAGA  
CAAGTTCTTCATCAACGGCTCCAACCTGGGAAGGCATCCTGGGGCTGGCCTAT  
GCTGAGATTGCCAGGCCTGACGACTCCCTGGAGCCTTTCTTTGACTCTCTGG  
TAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAGCTTTGTGGTGCTGG  
CTTCCCCCTCAACCAGTCTGAAGTGCTGGCCTCTGTGCGGAGGGAGCATGAT  
CATTGGAGGTATCGACCACTCGCTGTACACAGGCAGTCTCTGGTATACACCC  
ATCCGGCGGGAGTGATTATGAGGTGATCATTGTGCGGGTGGAGATCAAT  
GGACAGGATCTGAAAATGGAAGTGAAGGAGTACAACCTATGACAAGAGCATTG  
TGGACAGTGGCACCAACCAACCTTCGTTTGCCCAAGAAAGTGTGTTGAAGCTGC  
AGTCAAATCCATCAAGGCAGCCTCCTCCACGGAGAAGTTCCTGATGGTTTC  
TGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGCACCAACCCCTTGGAAC  
ATTTTCCAGTCATCTCACTCTACCTAATGGGTGAGGTTACCAACCAGTCCTT  
CCGCATCACCATCCTTCCGCAGCAATACCTGCGGCCAGTGGAAGATGTGGC  
CACGTCCCAAGACGACTGTTACAAGTTTGCCATCTCACAGTCATCCACGGGC  
ACTGTTATGGGAGCTGTTATCATGGAGGGCTTCTACGTTGTCTTTGATCGGG  
CCCGAAAACGAATTGGCTTTGCTGTCAGCGCTTGCCATGTGCACGATGAGTT  
CAGGACGGCAGCGGTGGAAGGCCCTTTGTACCTTGACATGGAAGACTG  
TGGCTACAACATTCCACAGACAGATGAGTCAACCCTCATGACCATAGCCTAT  
GTCATGGCTGCCATCTGCGCCCTCTTCATGCTGCCACTCTGCCTCATGGTGT  
GTCAGTGGCGCTGCCTCCGCTGCCTGCGCCAGCAGCATGATGACTTTGCTG  
ATGACATCTCCCTGCTGAAGTGAGGAGGCCCATGGGCAGAAGATAGAGATT  
CCCCTGGACCACACCTCCGTGGTTCACCTTTGGTCACAAGTAGGAGACACAGA  
TGGCACCTGTGGCCAGAGCACCTCAGGACCCTCCCCACCCACCAAATGCCT  
CTGCCTTGATGGAGAAGGAAAAGGCTGGCAAGGTGGGTTCAGGGACTGTA  
CCTGTAGGAAACAGAAAAGAGAAGAAAGAAAGCACTCTGCTGGCGGGAATAC  
TCTTGGTCACCTCAAATTTAAGTCGGGAAATTCTGCTGCTTGAAACTTCAGCC  
CTGAACCTTTGTCCACCATTCCTTTAAATTCTCCAACCCAAAGTATTCTTCTTT  
TCTTAGTTTCAGAAAGTACTGGCATCACACGCAGGTTACCTTGGCGTGTGTCC  
CTGTGGTACCCTGGCAGAGAAGAGACCAAGCTTGTTTCCCTGCTGGCCAAA  
GTCAGTAGGAGAGGATGCACAGTTTGCTATTTGCTTTAGAGACAGGGACTGT  
ATAACAAGCCTAACATTGGTGCAAAGATTGCCTCTTGAATT

**FIG. 1B**

MAQALPWLLLWMGAGVLP AHGTQH GIRLPLR SGLGGAPLGLRL  
PRETDEEPEEPGRRGSFVEMVDNLRGKSGQGYVEMTVGSPP  
QTLNILVDTGSSNFAVGAAPHPFLHRY YQRQLSSTYRDLRKG VY  
VPYTQGKWE GELGTDLV SIPHGPNVTVRANIAAITESDKFFINGS  
NWE GILGLAYAEIARPDDSLEPFFDSL VKQTHVPNLFSLQLCGAG  
FPLNQSEVLASVGGSMIIGGIDHSLYTGSLWYTPIRREWYYEVIIV  
RVEINGQDLKMDCKEYNYDKSIVDSGTTNLRLPKKVFEAAVKS IK  
AASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMGEVTN  
QSFRITILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIM  
EGFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDC  
GYNIPQTDESTLMTIAYVMAAICALFMLPLCLMVCQWRCLRCLR  
QQHDDFADDISLLK

**FIG. 2A**

ETDEEPEEPGRRGSFVEMVDNLRGKSGQGYYVEMTVGSPPT  
LNILVDTGSSNFAVGAAPHPFLHRYRQRQLSSTYRDLRKGVVYP  
YTQGWEGELGTDLVSIHPNVTVRANIAAITESDKFFINGSNW  
EGILGLAYAEIARPDDSLEPFFDSLQKQTHVPLFSLQLCGAGFP  
LNQSEVLASVGGSMIIGGIDHSLYTGSLWYTPIRREWYYEVIIVRV  
EINGQDLKMDCKEYNYDKSIVDSGTTNLRPLPKKVFEAAVKSIAA  
SSTEKFPDGFVLGEQLVCWQAGTTPWNIFPVISLYLMGEVTNQ  
SFRITILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIME  
GFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDC  
GYNIPQTDESTLMTIAYVMAAICALFMLPLCLMVCQWRCLRLR  
QQHDDFADDISLLK

**FIG. 2B**

**5/48**

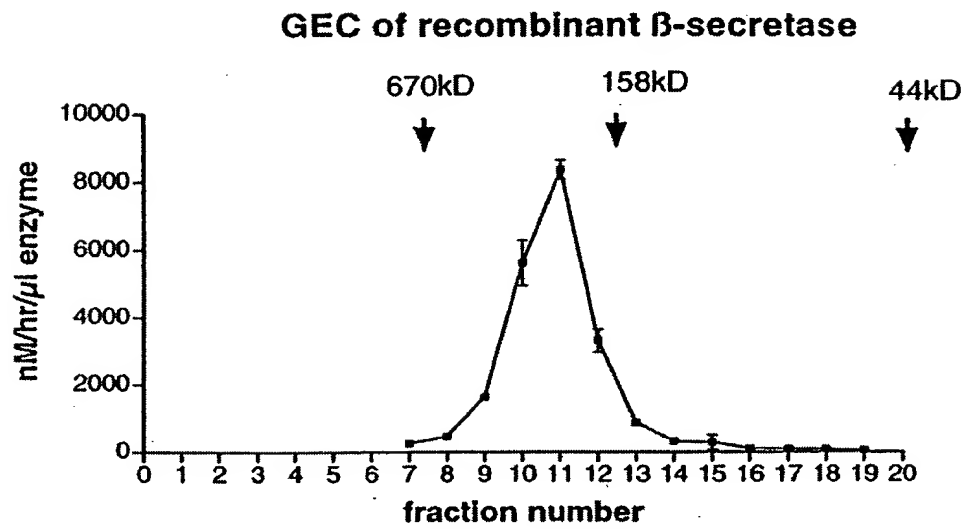
MAQALPWLLLWMGAGVLP AHGTQH GIRLPLRSG LGGAPLGLRL  
PRETDEEPEEPGRRGSFVEMVDNLRGKSGQGYYVEMTVGSPP  
QTLNILVDTGSSNFAVGAAPHPFLHRYYQRQLSSTYRDLRKGVY  
VPYTQGKWE GELGTDLV SIPHGPNVTVRANIAAITESDKFFINGS  
NWE GILGLAYAEIARPDDSLEPFFDSL VKQTHVPNLFSLQLCGAG  
FPLNQSEVLASVGGSMIIGGIDHSLYTGSLWYTPIRREWYYEVIIV  
RVEINGQDLKMDCKEYNYDKSIVDSGTTNLRLPKKVFEAAVKS IK  
AASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMGEVTN  
QSFRITILPQQYL RPVEDVATSQDDCYKFAISQSSTGTVMGAVIM  
EGFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDC  
GYNIPQTDEDYKDDDDK

**FIG. 3A**

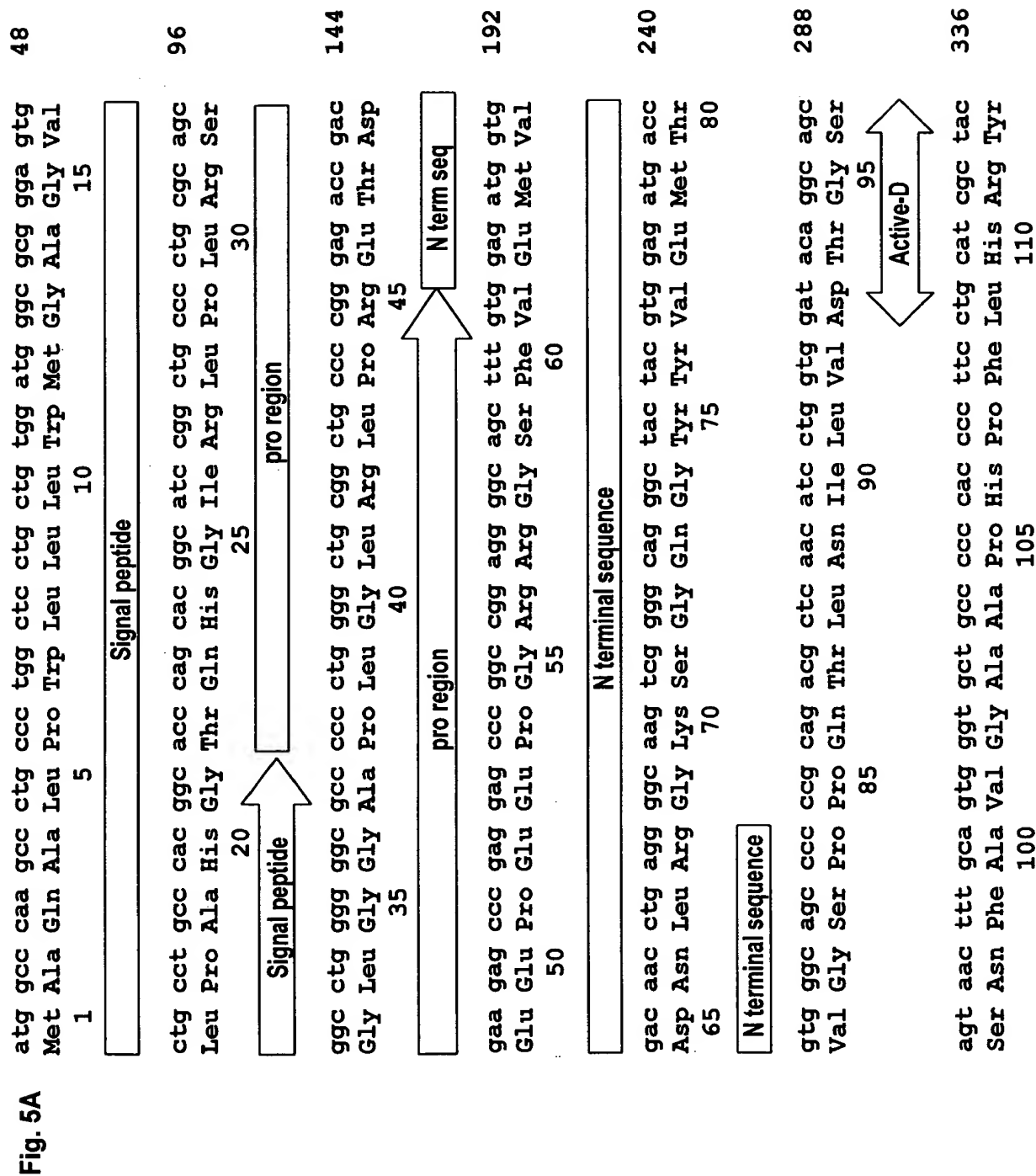
ETDEEPEEPGRRGSFVEMVDNLRGKSGQGYYVEMTVGSPPQT  
LNILVDTGSSNFAVGAAPHPFLHRYYQRQLSSTYRDLRKGVYVP  
YTQGKWE GELGTDLV SIPHGPNVTVRANIAAITESDKFFINGSNW  
EGILGLAYAEIARPDDSLEPFFDSL VKQTHVPNLFSLQLCGAGFP  
LNQSEVLASVGGSMIIGGIDHSLYTGSLWYTPIRREWYYEVIIVRV  
EINGQDLKMDCKEYNYDKSIVDSGTTNLRLPKKVFEAAVKSIAA  
SSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMGEVTNQ  
SFRITILPQQYL RPVEDVATSQDDCYKFAISQSSTGTVMGAVIME  
GFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDC  
GYNIPQTDEDYKDDDDK

**FIG. 3B**

6/48

**FIG. 4**

7/48



**8/48**

**Fig. 5B**

[illegible]



9/48

Fig. 5C

tct gaa gtg ctg gcc tct gtc gga ggg agc atg atc att gga ggt atc Ser Glu Val Leu Ala Ser Val Gly Gly Ser Met Ile Ile Gly Gly Ile 225 230 235 240	720
<b>N-gly</b>	
gac cac tcg ctg tac aca ggc agt ctc tgg tat aca ccc atc cgg cgg Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg 245 250 255	768
gag tgg tat tat gag gtg atc att gtg cgg gtg gag atc aat gga cag Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln 260 265 270	816
gat ctg aaa atg gac tgc aag gag tac aac tat gac aag agc att gtg Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val 275 280 285	864
gac agt ggc acc acc aac ctt cgt ttg ccc aag aaa gtg ttt gaa gct Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala 290 295 300	912
<b>Active-D</b>	
gca gtc aaa tcc atc aag gca gcc tcc tcc acg gag aag ttc cct gat Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp 305 310 315 320	960
ggt ttc tgg cta gga gag cag ctg gtg tgc tgg caa gca ggc acc acc Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr 325 330 335	1008

10/48

Fig. 5D

cct tgg aac att ttc cca gtc atc tca ctc tac cta atg ggt gag gtt Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val 340 345 350	1056
acc aac cag tcc ttc cgc atc acc atc ctt ccg cag caa tac ctg cgg Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg 355 360 365	1104
<div>N-glycos</div> cca gtg gaa gat gtg gcc acg tcc caa gac gac tgt tac aag ttt gcc Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys Tyr Lys Phe Ala 370 375 380	1152
atc tca cag tca tcc acg ggc act gtt atg gga gct gtt atc atg gag Ile Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu 385 390 395 400	1200
ggc ttc tac gtt gtc ttt gat cgg gcc cga aaa cga att ggc ttt gct Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala 405 410 415	1248
gtc agc gct tgc cat gtg cac gat gag ttc agg acg gca gcg gtg gaa Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu 420 425 430	1296
<div>Internal peptide sequence</div>	

11/48

Fig. 5E

ggc cct ttt gtc acc ttg gac atg gaa gac tgt ggc tac aac att cca	1344
Gly Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro	
435 440 445	
cag aca gat gag tca acc ctc atg acc ata gcc tat gtc atg gct gcc	1392
Gln Thr Asp Glu Ser Thr Leu Met Thr Ile Ala Tyr Val Met Ala Ala	
450 455 460	
Transmembrane	
atc tgc gcc ctc ttc atg ctg cca ctc tgc ctc atg gtg tgt cag tgg	1440
Ile Cys Ala Leu Phe Met Leu Pro Leu Cys Leu Met Val Cys Gln Trp	
465 470 475 480	
Transmembrane	
cgc tgc ctc cgc tgc ctg cgc cag cag cat gat gac ttt gct gat gac	1488
Arg Cys Leu Arg Cys Leu Arg Gln Gln His Asp Asp Phe Ala Asp Asp	
485 490 495	
atc tcc ctg ctg aag tga	1506
Ile Ser Leu Leu Lys	
500	

12/48

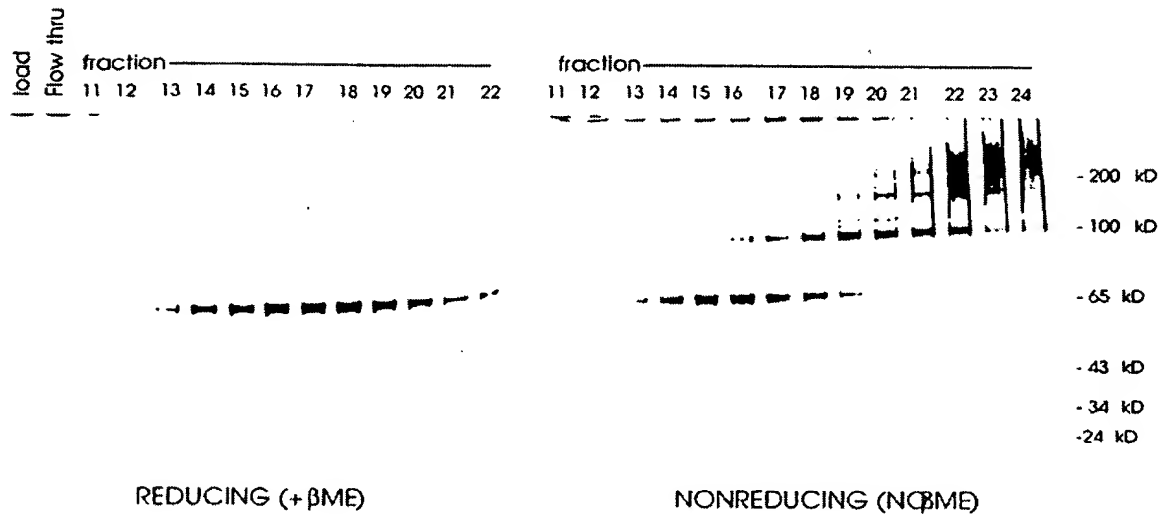


FIG. 6A

FIG. 6B

13/48

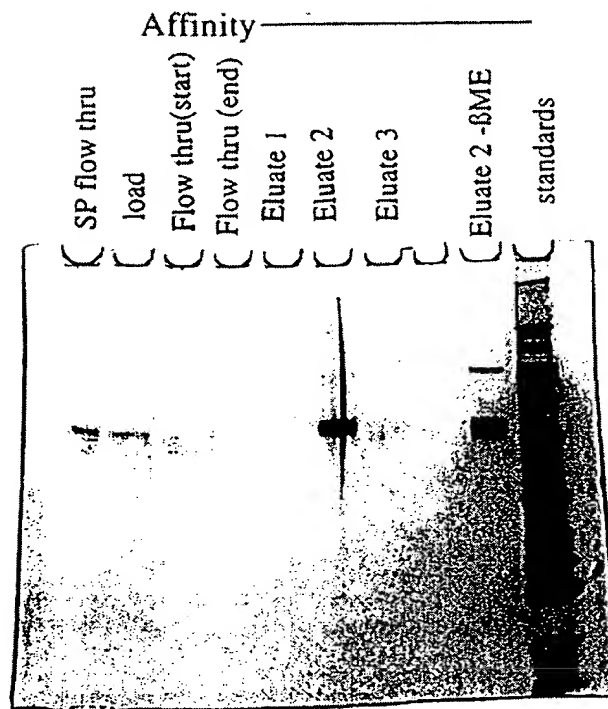


FIG. 7

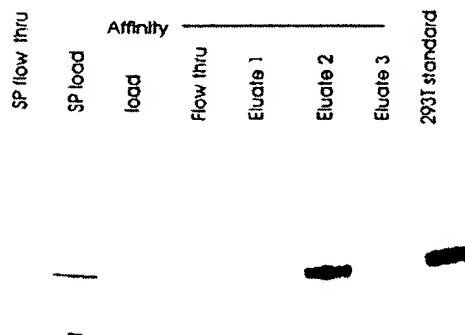


FIG. 8

14/48

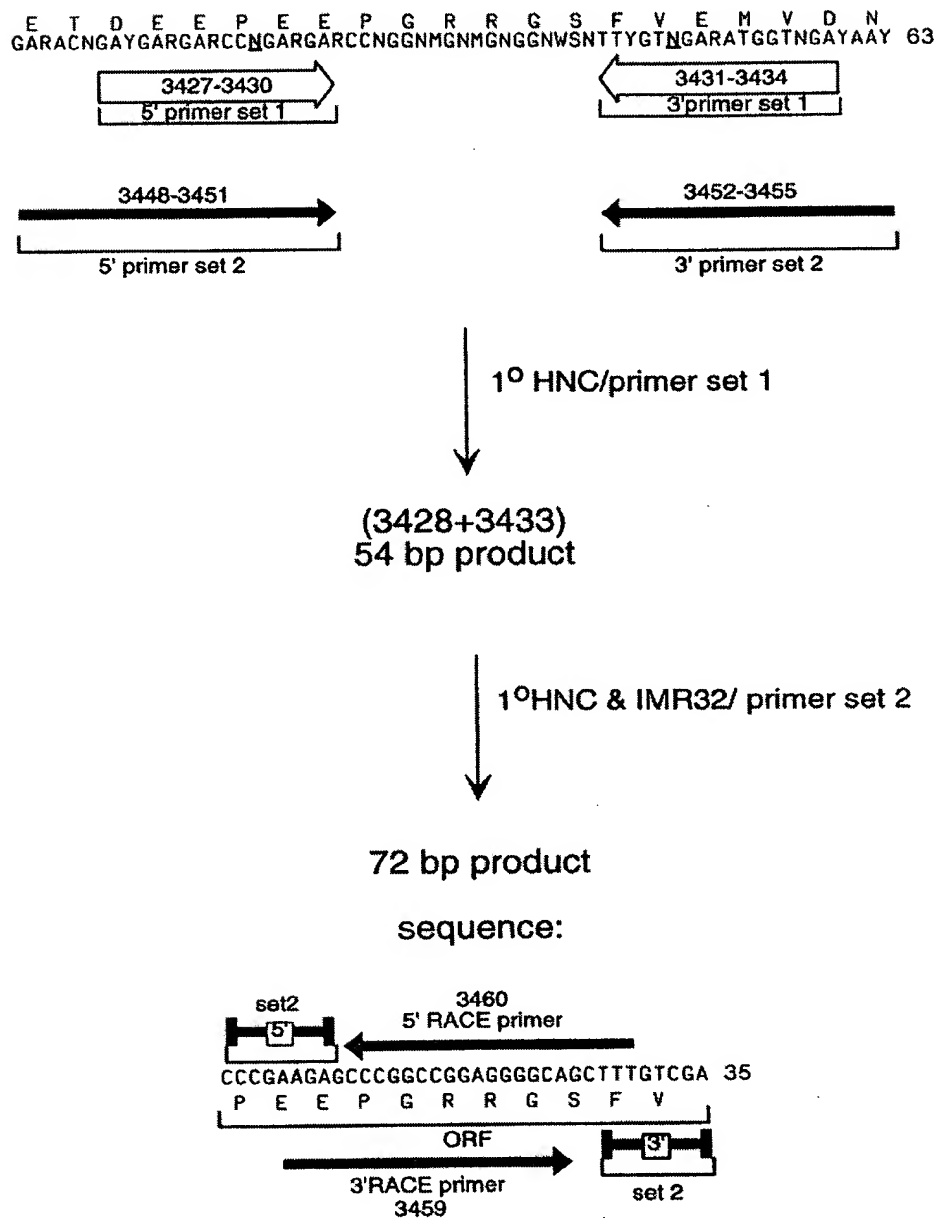


Fig. 9

15/48

	10	20	30	40	
Hump501prot	M A Q A L P W L L L W M G A G V L P A H G T Q H G I R L P L R S G L G G A P L G	40			
Musp501prot	M A P A L H W L L L W V G S G M L P A Q G T H L G I R L P L R S G L A G P P L G	40			
	50	60	70	80	
Hump501prot	L R L P R E T D E E P E E P G R R G S F V E M V D N L R G K S G Q G Y Y V E M T	80			
Musp501prot	L R L P R E T D E E S E E P G R R G S F V E M V D N L R G K S G Q G Y Y V E M T	80			
	90	100	110	120	
Hump501prot	V G S P P Q T L N I L V D T G S S N F A V G A A P H P F L H R Y Y Q R Q L S S T	120			
Musp501prot	V G S P P Q T L N I L V D T G S S N F A V G A A P H P F L H R Y Y Q R Q L S S T	120			
	130	140	150	160	
Hump501prot	Y R D L R K G V Y V P Y T Q G K W E G E L G T D L V S I P H G P N V T V R A N I	160			
Musp501prot	Y R D L R K G V Y V P Y T Q G K W E G E L G T D L V S I P H G P N V T V R A N I	160			
	170	180	190	200	
Hump501prot	A A I T E S D K F F I N G S N W E G I L G L A Y A E I A R P D D S L E P F F D S	200			
Musp501prot	A A I T E S D K F F I N G S N W E G I L G L A Y A E I A R P D D S L E P F F D S	200			
	210	220	230	240	
Hump501prot	L V K Q T H Y P N L F S L Q L C G A G F P L N Q S E V L A S V G G S M I I G G I	240			
Musp501prot	L V K Q T H I P N I F S L Q L C G A G F P L N Q T E A L A S V G G S M I I G G I	240			
	250	260	270	280	
Hump501prot	D H S L Y T G S L W Y T P I R R E W Y Y E V I I V R V E I N G Q D L K M D C K E	280			
Musp501prot	D H S L Y T G S L W Y T P I R R E W Y Y E V I I V R V E I N G Q D L K M D C K E	280			
	290	300	310	320	
Hump501prot	Y N Y D K S I V D S G T T N L R L P K K V F E A A V K S I K A A S S T E K F P D	320			
Musp501prot	Y N Y D K S I V D S G T T N L R L P K K V F E A A V K S I K A A S S T E K F P D	320			
	330	340	350	360	
Hump501prot	G F W L G E Q L V C W Q A G T T P W N I F P V I S L Y L M G E V T N Q S F R I T	360			
Musp501prot	G F W L G E Q L V C W Q A G T T P W N I F P V I S L Y L M G E V T N Q S F R I T	360			
	370	380	390	400	
Hump501prot	I L P Q Q Y L R P V E D V A T S Q D D C Y K F A I S Q S S T G T V M G A V I M E	400			
Musp501prot	I L P Q Q Y L R P V E D V A T S Q D D C Y K F A V S Q S S T G T V M G A V I M E	400			
	410	420	430	440	
Hump501prot	G F Y V V F D R A R K R I G F A V S A C H V H D E F R T A A V E G P F V T L D M	440			
Musp501prot	G F Y V V F D R A R K R I G F A V S A C H V H D E F R T A A V E G P F V T A D M	440			
	450	460	470	480	
Hump501prot	E D C G Y N I P Q T D E S T L M T I A Y V M A A I C A L F M L P L C L M V C Q W	480			
Musp501prot	E D C G Y N I P Q T D E S T L M T I A Y V M A A I C A L F M L P L C L M V C Q W	480			
	490	500			
Hump501prot	R C L R C L R Q Q H D D F A D D I S L L K				501
Musp501prot	R C L R C L R H Q H D D F G D D I S L L K				501

FIG 10

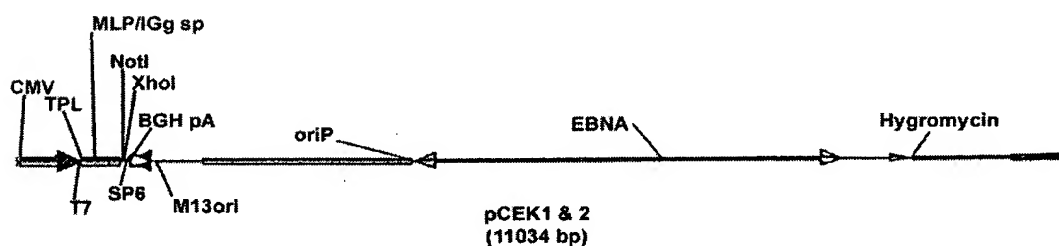
FIG. 10

501

501

**16/48**

CTGTTGGGCTCGCGGTGAGGACAACTCTTCGCGGTCTTTCCAGTACTCT  
 TGGATCGGAAACCGTCGGCCTCCGAACGGTACTCCGCCACCGAGGGACCT  
 GAGCGAGTCCGCATCGACCGGATCGGAAAACCTCTCGACTGTTGGGGTGAG  
 TACTCCCTCTCAAAGCGGGCATGACTTCTGCGCTAAGATTGTCAGTTTCC  
 AAAACGAGGAGGATTTGATATTCACCTGGCCCGCGGTGATGCCTTTGAGG  
 GTGGCCGCGTCCATCTGGTCAGAAAAGACAATCTTTTTGTTGTCAAGCTTG  
 AGGTGTGGCAGGCTTGAGATCTGGCCATACACTTGAGTGACAATGACATCC  
 ACTTTGCCCTTTCTCTCCACAGGTGTCCACTCCCAGGTCCAAGTGCAGGTCTG  
 ACTCTAGACCC

**FIG. 11A****FIG. 11B**



17/48

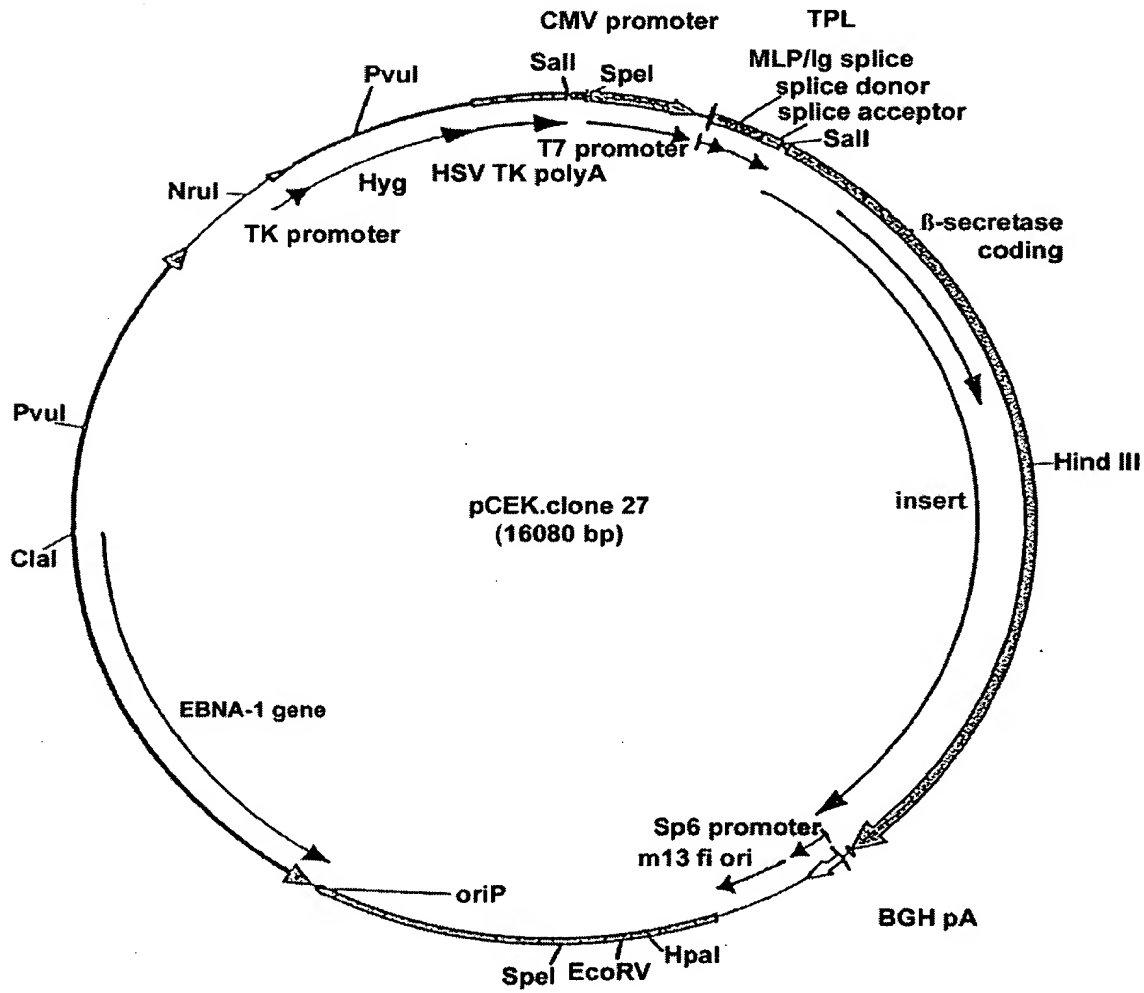


FIG. 12

18/48

### Figure 13A

ttctcatgtt tgacagctta tcatcgaga tccgggcaac gttgttgcac tgctgcaggc 60  
gcagaactgg taggtatgga agatccgatg tacggggccag atatacgcgt tgacattgat 120  
SpeI  
tattgactag ttattaatag taatcaatta cggggtcatt agttcatagc ccatatatgg 180  
agttccgcgt tacataactt acggtaaatg gccgccttg ctgaccgccc aacgaccccc 240  
gcccattgac gtcaataatg acgtatgttc ccatagtaac gccaataggg actttccatt 300  
gacgtcaatg ggtggactat ttacggtaaa ctgccactt ggcagtacat caagtgtatc 360  
atatgccaag tacgccccct attgacgtca atgacggtaa atggcccgcc tggcattatg 420  
ccagtacat gaccttatgg gactttccta ctgggcagta catctacgta ttagtcatcg 480  
ctattaccat ggtgatgcgg ttttggcagt acatcaatgg gcgtggatag cggtttgact 540  
cacgggggatt tcaaagtctc caccaccattg acgtcaatgg gagtttgttt tggcaccaaa 600  
atcaacggga ctttccaaaa tgtcgtaaca actccgcccc attgacgcaa atgggcggtgta 660  
ggcgtgtacg gtgggaggtc tataaagca gagctctctg gctaaactaga gaaccactg 720  
cttactggct tatcgaaatt aatacgactc actataggga gacccaagct ctgttgggct 780

19/48

Figure 13B

cgcggttgag gacaaactct tcgcggtctt tccagtactc ttggatcgga aaccgcgtcg 840  


---

 cctccgaacg gtactccgcc accgaggagc ctgagcgagt ccgcatcgac cggatcggaa 900  
 splice donor  
 aacctctcga ctgttgggggt gagtactccc tctcaaaagc gggcatgact tctgcgctaa 960  


---

 gattgtcagt ttccaaaaac gaggaggatt tgatatcac ctggcccgcg gtgatgcctt 1020  
 tgaggggtggc cgcgtccatc tggtcagaaa agacaatctt tttgttgtca agcttgaggt 1080  


---

 gtggcaggct tgagatctgg ccatacactt gagtgaacaat gacatccact ttgcctttct 1140  
 splice acceptor SalI  
 ctccacaggt gtccactccc aggtccaact gcaggctcgac tctagaccgc gggaattctg 1200  
 cagatatcca tcacactggc cgcactcgtc ccagcccgc cgggagctg cgagccgcga 1260  
 gctggattat ggtggcctga gcagccaacg cagccgcagg agcccggagc ccttgcccct 1320  
 gcccgcgccg ccgcccgcg gggggaccag ggaagccgc accggcccgc catgcccgcc 1380  


---

 cctcccagcc ccgccgggag cccgcgcccg ctgccaggc tggccgcccgc cgtgccgatg 1440  
 tagcgggctc cggatcccag cctctcccct gctcccgtgc tctgaggatc tcccctgacc 1500  


---

 gctctccaca gcccggaccc gggggctggc ccaggggccct gcaggcccctg gcgtcctgat 1560  


---

 gcccccaagc tccctctcct gagaagccac cagcaccacc cagacttggg ggcaggcgcc 1620

20/48

Figure 13C

1677	agggacggac	gtgggccagt	gcgagcccag	agggcccgaa	ggccggggcc	cacc	atg	Met
								1
1725	gcc caa	gcc ctg	ccc tgg	ctc ctg	ctg tgg	atg ggc	gga gtg	ctg
	Ala Gln	Ala Leu	Pro Trp	Leu Leu	Trp Met	Gly Ala	Gly Val	Leu
		5		10			15	
1773	cct gcc	cac ggc	acc cag	cac ggc	atc cgg	ctg ccc	ctg cgc	agc ggc
	Pro Ala	His Gly	Thr Gln	His Gly	Ile Arg	Leu Pro	Leu Arg	Ser Gly
		20		25		30		
1821	ctg ggg	ggc gcc	ccc ctg	ggg ctg	cgg ctg	ccc cgg	gag acc	gac gaa
	Leu Gly	Gly Ala	Pro Leu	Gly Leu	Arg Leu	Pro Arg	Glu Thr	Asp Glu
		35		40		45		
1869	gag ccc	gag gag	ccc ggc	cgg agg	ggc agc	ttt gtg	gag atg	gtg gac
	Glu Pro	Glu Glu	Pro Gly	Arg Arg	Gly Ser	Phe Val	Glu Met	Val Asp
		50		55		60		65
1917	aac ctg	agg ggc	aag tcg	ggg cag	ggc tac	tac gtg	gag atg	acc gtg
	Asn Leu	Arg Gly	Lys Ser	Gly Gln	Gly Tyr	Tyr Val	Glu Met	Thr Val
		70		75		80		
1965	ggc agc	ccc ccg	cag acg	ctc aac	atc ctg	gtg gat	aca ggc	agc agt
	Gly Ser	Pro Pro	Gln Thr	Leu Asn	Ile Leu	Val Asp	Thr Gly	Ser Ser
		85		90		95		

21/48

Figure 13D

2013	aac ttt gca gtg ggt gct gcc ccc cac ccc ttc ctg cat cgc tac tac	100	105	110	115
	Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr Tyr				
2061	cag agg cag ctg tcc agc aca tac cgg gac ctc cgg aag ggt gtg tat	120	125		
	Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val Tyr				
2109	gtg ccc tac acc cag gcc aag tgg gaa ggg gag ctg ggc acc gac ctg	135	140	145	
	Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp Leu				
2157	gta agc atc ccc cat gcc ccc aac gtc act gtg cgt gcc aac att gct	150	155	160	
	Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg Ala Asn Ile Ala				
2205	gcc atc act gaa tca gac aag ttc atc atc aac ggc tcc aac tgg gaa	165	170	175	
	Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp Glu				
2253	ggc atc ctg ggg ctg gcc tat gct gag att gcc agg cct gac gac tcc	180	185	190	
	Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro Asp Asp Ser				
2301	ctg gag cct ttc ttt gac tct ctg gta aag cag acc cac gtt ccc aac	195	200	205	
	Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr His Val Pro Asn				

22/48

Figure 13E

ctc ttc tcc ctg cag ctt tgt ggt gct ggc ttc ccc ctc aac cag tct	2349
Leu Phe Ser Leu Gln Leu Cys Gly Ala Gly Phe Pro Leu Asn Gln Ser	
210 215 220 225	
<hr/>	
gaa gtg ctg gcc tct gtc gga ggg agc atg atc att gga ggt atc gac	2397
Glu Val Leu Ala Ser Val Gly Gly Ser Met Ile Ile Gly Gly Ile Asp	
230 235 240	
<hr/>	
cac tcg ctg tac aca ggc agt ctc tgg tat aca ccc atc cgg cgg gag	2445
His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg Glu	
245 250 255	
<hr/>	
tgg tat tat gag gtc atc att gtg cgg gtg gag atc aat gga cag gat	2493
Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln Asp	
260 265 270	
<hr/>	
ctg aaa atg gac tgc aag gag tac aac tat gac aag agc att gtg gac	2541
Leu Lys Met Asp Cys Lys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val Asp	
275 280 285	
<hr/>	
agt ggc acc acc aac ctt cgt ttg ccc aag aaa gtg ttt gaa gct gca	2589
Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala Ala	
290 295 300 305	
<hr/>	
gtc aaa tcc atc aag gca gcc tcc tcc acg gag aag ttc cct gat ggt	2637
Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp Gly	
310 315 320	

23/48

Figure 13F

2685	ttc tgg cta gga gag cag ctg gtg tgc tgg caa gca ggc acc acc cct	325	330	335
	Phe Trp Leu Gly Gln Leu Val Cys Trp Gln Ala Gly Thr Thr Pro			
2733	tgg aac att ttc cca gtc atc tca ctc tac cta atg ggt gag gtt acc	340	345	350
	Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val Thr			
2781	aac cag tcc ttc cgc atc acc atc ctt ccg cag caa tac ctg cgg cca	355	360	365
	Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg Pro			
2829	gtg gaa gat gtg gcc acg tcc caa gac gac tgt tac aag ttt gcc atc	370	375	380
	Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys Tyr Lys Phe Ala Ile			
2877	tca cag tca tcc acg ggc act gtt atg gga gct gtt atc atg gag ggc	390	395	400
	Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu Gly			
2925	ttc tac gtt gtc ttt gat cgg gcc cga aaa cga att ggc ttt gct gtc	405	410	415
	Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala Val			
2973	agc gct tgc cat gtg cac gat gag ttc agg acg gca gcg gtg gaa ggc	420	425	430
	Ser Ala Cys His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu Gly			

24/48

Figure 13G

cct ttt gtc acc ttg gac atg gaa gac tgt ggc tac aac att cca cag	3021
Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro Gln	
435 440 445	
<hr/>	
aca gat gag tca acc ctc atg acc ata gcc tat gtc atg gct gcc atc	3069
Thr Asp Glu Ser Thr Leu Met Thr Ile Ala Tyr Val Met Ala Ala Ile	
450 455 460 465	
<hr/>	
tgc gcc ctc ttc atg ctg cca ctc tgc ctc atg gtg tgt cag tgg cgc	3117
Cys Ala Leu Phe Met Leu Pro Leu Cys Leu Met Val Cys Gln Trp Arg	
470 475 480	
<hr/>	
tgc ctc cgc tgc ctg cgc cag cag cat gat gac ttt gct gat gac atc	3165
Cys Leu Arg Cys Leu Arg Gln Gln His Asp Asp Phe Ala Asp Asp Ile	
485 490 495	
<hr/>	
tcc ctg ctg aag tga ggaggcccat ggcagaaga tagagattcc cctggaccac	3220
Ser Leu Leu Lys	
500	
<hr/>	
acctccgtgg ttcaacttgg tcacaagtag gagacacaga tggcacctgt ggccagagca	3280
<hr/>	
cctcaggacc ctcccaccc accaaatgcc tctgccttga tggagaagga aaaggctggc	3340
<hr/>	
aagggtgggtt ccagggactg tacctgtagg aaacagaaaa gagaagaaag aagcactctg	3400
<hr/>	
ctggcgggaa tactcttggt cacctcaaat ttaagtcggg aaattctgct gcttgaaact	3460
<hr/>	



**25/48**

## Figure 13H

**tcagccctga acctttgtcc accattcctt taaattctcc aacccaaagt attcttcttt 3520**

tcttagtttc agaagtactg gcatcacag caggttacct tggcgtgtgt cctgtggtta 3580

# HindiII

ccctggcaga gaagagacca agcttgtttc cctgctggcc aaagtcagta ggagaggatg 3640

**cacagtttgc tatttgctt agagacaggg actgtataaa caagcctaac attggtgcaa 3700**

**agattgctc ttgaattaa aaaaaaact agattgacta ttatacaa tggggcggc 3760**

**tggaaagagg agaaggagag ggagtacaaa gacaggggaat agtgggatca aagctaggaa 3820**

aggcagaac acaaccactc accagtccta gtttagacc tcattccaa gatagcatcc 3880

catctcagaa gatgggtgtt gtttcaatg ttttctttc tgtggttgca gcctgaccaa 3940

aagtgagatg ggaagggtt atctagccaa agagctcttt ttagctctc ttaaatgaag 4000

**tgcccactaa gaagttccac ttaacacatg aatttctgcc atattaattt cattgtctct 4060**

atctgaacca cccttattc tacatatgat aggcagcact gaaatatacct aacccctaa 4120

gctccagggtg ccctgtggga gagcaactgg actatagcag ggctgggctc tgtcttctcg 4180

gtcataggct cactcttcc ccaaatctt cctctggagc ttgcagcca agtgctaaa 4240

aggaataggt aggagacctc ttctatctaa tccttaaag cataatgttg aacattcatt 4300

Figure 13I

caacagctga tgcctataa cccctgcctg gatttcttcc tattaggcta taagaagtag 4360  
caagatcttt acataattca gagtggtttc attgccttcc taccctctct aatggcccct 4420  
ccatttattt gactaaagca tcacacagtg gcactagcat tataccaaga gtatgagaaa 4480  
tacagtgcct tatggctcta acattactgc cttcagtatc aaggctgcct ggagaaaagga 4540  
tggcagcctc agggcttctt tatgtcctcc accacaagag ctcttgatg aaggtcatct 4600  
ttttcccta tcctgttctt cccctccccg ctctaattgg tacgtgggta cccaggctgg 4660  
ttcttgggct aggtagtggg gaccaagtcc attacctccc tatcagttct agcatagtaa 4720  
actacgggtac cagtgttagt ggaagagct gggttttctt agtataccca ctgcctccta 4780  
ctcctacctg gtcaacccgc tgcttccagg tatgggacct gctaagtgtg gaattacctg 4840  
ataagggaga gggaaataca aggagggcct ctggtgttcc tggcctcagc cagctgcccc 4900  
caagccataa accaataaaa caagaatact gagtcaagttt ttatatctggg ttctcttcat 4960  
tcccactgca cttgggtgctg ctttggctga ctgggaacac ccataacta cagagtctga 5020  
caggaagact ggagactgtc cacttctagc tcggaaactta ctgtgtaaat aaactttcag 5080  
aactgctacc atgaagtga aatgccacat ttgtcttat aattctacc catgttggga 5140

## 27/48

Figure 13J

aaaactggct ttttccagc cttttccagg gcataaaact caacccttc gatagcaagt 5200

cccatcagcc tattatttt ttaagaaaa ctgcacttg ttttctttt tacagttact 5260

tccttctgc cccaaaatta taaactctaa gtgtaaaaaa aagtcttaac aacagcttct 5320

tgcttgtaaa aatatgtatt atacatctgt attttaaat tctgctcctg aaaaatgact 5380

gtccattct cactcactg catttggggc ctttccatt ggtctgcatg tcttttatca 5440

ttgcaggcca gtggacagag ggagaaggga gaacaggggt cgccaacact tgttgtgctt 5500

tctgactgat cctgaacaag aaagagtaac actgaggcgc tcgctcccat gcacaactct 5560

ccaaaacact tatctctctg caagagtggg ctttccgggt ctttactggg aagcagttaa 5620

gccccctct cacccttcc tttttcttt ctttactcct ttggcttcaa aggatttgg 5680

aaaagaaaca atatgcttta cactcathtt caatttctaa atttgcaagg gatactgaaa 5740

aatacggcag gtggcctaag gctgctgtaa agtgagggg agaggaaatc ttaagattac 5800

aagataaaaa acgaatcccc taaacaaaaa gaacaataga actggtcttc cattttgcc 5860

ccttctctgt tcatgacagc tactaacctg gagacagtaa catttcatta accaaaagaa 5920

gtgggtcacc tgacctctga agagctgagt actcaggcca ctccaatcac cctacaagat 5980

28/48

Figure 13K

gccaaaggagg tccaggaag tccagctcct taaactgacg ctagtcaata aacctgggca 6040  
agtaggcaa gagaaatgag gaagaatcca tctgtgaggt gacaggcacg gatgaagac 6100  
aaagacggaa aagagtatca aaggcagaaa ggagatcatt tagttgggtc tgaaaggaaa 6160  
agtntttgct atccgacatg tactgctagt wcctgtaagc atttaggtc ccagaatgga 6220  
aaaaaaaaatc aagctatngg ttatataata atgnnnnnnn nnnnnnnnn nntcgagcat 6280  
gcatctagag ggcctattc tatagtgtca cctaaatgct agagctgct gatcagcctc 6340  
gactgtgcct tctagttgcc agccatctgt tgtttgccc tccccgtgc ctctcttgac 6400  
cctggaaggt gccactccca ctgtcccttc ctaataaaat gaggaattg catcgcatg 6460  
tctgagtagg tgtcattcta ttctgggggg tggggtgggg caggacagca agggggagga 6520  
ttgggaagac aatagcaggc atgctgggga tgcggtgggc tctatggctt ctgaggcga 6580  
aagaaccagc tggggctcta gggggatatcc ccacgcgcc tgtagcggcg cattaagcgc 6640  
ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt gccagcgccc tagcggccc 6700  
tcctttcgct ttcttcctt cttttctgc cacttcgcc ggctttccc gtcaagctct 6760  
aaatcggggc atccctttag ggttcgatt tagtgcttta cggcacctcg accccaaaa 6820

29/48

Figure 13L

acttgattag ggtgatggtt cacgtagtgg gccatcgccc tgatagcgg ttttcgccc 6880  
tttgacgttg gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact 6940  
caaccctatc tcggtctatt cttttgattt ataagggatt ttggggattt cggcctattg 7000  
 gttaaaaaat gagctgattt acaaaaaatt taacgcgaat tctagagccc cgccgcgga 7060  
 cgaactaac ctgactacgg catctctgcc ccttcttcgc ggggcagtgc atgtaatccc 7120  
 ttcagttggt tggtaacaact tgccaaactgg gccctgttcc acatgtgaca cgggggggga 7180  
 ccaaacacaa aggggttctc tgactgtagt tgacatcctt ataaatggat gtgcacattt 7240  
 gccaaactg agtggcttcc atcctggagc agactttgca gtctgtggac tgcaaacacaa 7300  
 cattgccttt atgtgtaact cttggctgaa gctcttacac caatgctggg ggacatgtac 7360  
 ctcccagggg cccaggaaga ctacgggagg ctacaccaac gtcaatcaga ggggcctgtg 7420  
 tagctaccga taagcgacc ctcaagaggg cattagcaat agtgtttata agccccctt 7480  
 gttaacccta aacgggtagc atatgcttcc cgggtagtag tatatactat ccagactaac 7540  
 cctaattcaa tagcatatgt taccacaagg gaagcatatg ctatcgaatt agggtagta 7600  
 aaagggtcct aaggaacagc gatattctccc accccatgag ctgtcacggg tttatttaca 7660

↑  
 HpaI  
 EcoRV

30/48

Figure 13M

tggggtcagg attccacgag ggtagtgaac cattttagtc acaagggcag tggctgaaga 7720  
tcaaggagcg ggcagtgaac tctcctgaat ctctgcctgc ttcttcattc tccttcgttt 7780  
agctaataga ataactgctg agtttgtgaac agtaagggtg atgtgagggtg ctcgaaaaaca 7840  
aggtttcagg tgacgcccc agaataaaat ttggacgggg ggttcagtggtg tggcatgtg 7900  
ctatgacacc aatataaacc tcacaaaacc ctggggcaat aaatactagt gtaggaatga 7960  
aacattctga atatctttaa caatagaaat ccatgggggtg gggacaagcc gtaaagactg 8020  
gatgtccatc tcacacgaat ttatggctat gggcaacaca taatccctagt gcaatatgat 8080  
actgggggta ttaagatgtg tcccaggcag ggaccaagac aggtgaacca tgttgttaca 8140  
ctctatttgt aacaagggga aagagagtgg acgccgacag cagcggactc cactggttgt 8200  
ctctaacacc cccgaaaaat aaacgggggt ccacgccaat gggggccata aacaaaagaca 8260  
agtggccact cttttttttg aaatttgtga gtgggggcac gcgtcagccc ccacacgccg 8320  
ccctgcggtt ttggactgta aaataagggt gtaataactt ggctgattgt aaccccgcta 8380  
accactgcgg tcaaaaccact tgcccacaaa accactaatg gcaccccggt gaatacctgc 8440  
ataagtaggt gggcggggcca agataggggc gcgattgctg cgatctggag gacaaaattac 8500

31/48

Figure 13N

acacactgc gcctgagcg caagcacagg gttgttggtc ctcatattca cgaggctgct 8560  
gagagcacgg tgggctaag ttgccatggg tagcatatac taccacaata tctggatagc 8620  
atatgctatc ctaatctata tctgggtagc ataggctatc ctaatctata tctgggtagc 8680  
atatgctatc ctaatctata tctgggtagt atatgctatc ctaattata tctgggtagc 8740  
ataggctatc ctaatctata tctgggtagc atatgctatc ctaatctata tctgggtagt 8800  
atatgctatc ctaatctgta tccgggtagc atatgctatc ctaatagaga ttaggtagt 8860  
atatgctatc ctaatttata tctgggtagc atatactacc caaatatctg gatagcatat 8920  
gctatcctaa tctatatctg ggtagcatat gctatcctaa tctatatctg ggtagcatag 8980  
gctatcctaa tctatatctg ggtagcatat gctatcctaa tctatatctg ggtagcatat 9040  
gctatcctaa tttatatctg ggtagcatag gctatcctaa tctatatctg ggtagcatat 9100  
gctatcctaa tctatatctg ggtagtatat gctatcctaa tctgtatccg ggtagcatat 9160  
gctatcctca tgcatatata gtcagcatat gataccagc agtagagtgg gagtgcctatc 9220  
ctttgcatat gccgccacct cccaagggg cgtgaattt cgctgcttgt ccttttcttg 9280  
oriP  
catgctgggt gctcccatc ttaggtgaat ttaaggaggc caggctaaag ccgtcgcctg 9340

32/48

Figure 130

tctgattgct caccaggtaa atgtcgctaa tgttttccaa cgcgagaagg tgttgagcgc 9400

ggagctgagt gacgtgacaa catgggtatg cccaattgcc ccatgttggg aggacgaaaa 9460

tggtgacaag acagatggcc agaaatacac caacagcacg catgatgtct actgggggatt 9520

tattctttag tgcgggggaa tacacggcctt ttaatacgat tgagggcgctc tcctaacaag 9580

ttacatcact cctgccccttc ctacccctca tctccatcac ctcttcac tccgtcatct 9640

ccgtcatcac cctccgcggc agccccttc accataggtg gaaaccaggg aggcaaatct 9700

actccatcgt caaagctgca cacagtcacc ctgatatgac aggtaggagc gggcctttgtc 9760

ataacaaggc ccttaatgc atcctcaaa acctcagcaa atatatgagt ttgtaaaaag 9820

acatgaaat aacagacaat ggactccctt agcggggccag gttgtgggcc gggtcaggcg 9880

gccattccaa aggggagacg actcaatggt gtaagacgac attgtggaat agcaagggca 9940

gttcctcgcc ttaggttgta aaggaggctc ttactacctc catatacgaa cacaccggcg 10000

acccaagttc cttcgtcggc agtcctttct acgtgactcc tagccaggag agctcttaaa 10060

ccttctgcaa tgttctcaa ttctgggttg gaacctcctt gaccacgatg cttccaaac 10120

caccctcctt ttttgccct gcctccatca cctgacccc ggggtccagt gcttgggcct 10180



**33/48****Figure 13P**

tctcctgggt catctgcggg gccctgctct atcgctcccc ggggcacgtc aggctcacca 10240  
tctgggccac cttcttggtg gtattcaaaa taatcggctt cccctacagg gtggaaaaat 10300  
ggccttctac ctggaggggg cctgcgcggt ggagacccgg atgatgatga ctgactactg 10360  
ggactcctgg gcctcttttc tcacagtcca cgacctctcc cctgggctct ttacgactt 10420  
ccccccctgg ctctttcacg tcctctaccc cggcggcctc cactacctcc tcgaccccg 10480  
cctccactac ctctcggacc cgggctcca ctgcctctc gacccggcc tcacctcct 10540  
gctcctgccc ctctgctcc tgccccctct cctgctctg cccctcctgc cctcctgct 10600  
cctgccccctc ctgccccctcc tgctcctgccc cctcctgccc ctctgctcc tgccccctc 10660  
gccccctcctc ctgctcctgc cctcctgccc cctcctcctg ctctgcccc tctgcccc 10720  
cctgctcctg cccctcctgc cctcctgct cctgccccctc ctgccccctc tgctcctgcc 10780  
cctcctgctc ctgccccctcc tgctcctgccc cctcctgctc ctgccccctc tgccccctc 10840  
gccccctcctc ctgctcctgc cctcctgct cctgccccctc ctgccccctc tgccccctc 10900  
gctcctgccc ctctcctgc tctgccccct cctgccccctc ctgccccctc tctgctcct 10960  
gccccctcctg cccctcctcc tgctcctgcc cctcctcctg ctctgcccc tctgcccc 11020

34/48

Figure 13Q

cctgcccctc ctctctgtctc tgcccctcct gccctctctc ctgtctctgc cctctctctc 11080  
gtctctgccc ctcttgcccc tcttgcccct cctctctgtc ctgcccctcc tctgtctctc 11140  
gcccctcctg cccctctctgc cctctctgccc cctctctctg ctcttgcccc tctctctgtc 11200  
cctgcccctc ctgtctctgc cctctcccgt cctgtctctg ctctgtgtcc accgtgggtc 11260  
cctttgcagc caatgcaact tggacgtttt tggggtctcc ggacaccatc tctatgtctt 11320  
ggccctgac ctgagccgcc cggggctctc ggtcttcgc ctctctgtcc tctctctctt 11380  
ccccgtctc gtccatggtt atcaccccct ctcttttgag gtccactgcc gccggagcct 11440  
tctggtccag atgtgtctcc ctctctctct aggcatttc caggtctctgt acctggcccc 11500  
tcgtcagaca tgattcacac taaaagagat caatagacat cttattaga cgacgtcag 11560  
tgaatacagg gagtgcagac tcttgcccc tccaacagcc ccccaccct catccccttc 11620  
atggtcgtcgc tcagacagat ccagggtctga aaattcccca tctccgaac catctctgtc 11680  
ctcatcacca attactcgca gcccggaaaa ctcccgtga acatcctcaa gatttgcgtc 11740  
ctgagcctca agccaggcct caaattctc gtccccctt ttgctggacg gtaggatgg 11800  
ggattctcgg gaccctctc ctctctcttc aaggtcacca gacagagatg ctactggggc 11860

35/48

Figure 13R

ClaI

aacggaagaa aagctgggtg cggcctgtga ggatcagctt atcgatgata agctgtcaaa 11920  
 catgagaatt cttgaagacg aaagggcctc gtgatacgcc tatatttata ggttaatgtc 11980  
 atgataataa tggttttctta gacgtcaggt ggcacttttc ggggaaatgt gcgcggaacc 12040  
 cctatttggt tatttttcta aatacatcca aatatgtatc cgctcatgag acaataaacc 12100  
 tgataaatgc ttcaataata ttgaaaaagg aagagtatga gtattcaaca ttcccggtgc 12160  
 gccctattc ccttttttgc ggcattttgc ctccctgttt ttgctcacc agaaacgctg 12220  
 gtgaaaagtaa aagatgctga agatcagttg ggtgcacgag tgggttacat cgaactggat 12280  
 ctcaacagcg gtaagatcct tgagagtttt cgccccgaag aacgtttttcc aatgatgagc 12340  
 acttttaaa ttctgctatg tggcgcggtat ttatcccgtg ttgacgccgg gcaagagcaa 12400  
 ctcgggtcgcc gcatacacta ttctcagaat gacttggttg agtactcacc agtcacagaa 12460  
 aagcatctta cggatggcat gacagtaaga gaattatgca gtgctgccat aacctagat 12520  
 gataaactg cggccaaactt acttctgaca acgatcggag gaccgaagga gctaaccgct 12580  
 tttttgcaca acatgggga tcatgtaact cgccttgatc gttgggaacc ggagctgaat 12640  
 gaagccatac caaacgacga gcgtgacacc acgatgcctg cagcaatggc aacaacgttg 12700

PvuI

Figure 13S

cgcaaatat taactggcg aactactact ctactctccc ggcaacaatt aatagactgg 12760  
atggaggcgg ataaaagtgc aggaccatt ctgcgctcgg cccttccggc tggctggttt 12820  
attgctgata aatctggagc cggtgagcgt gggctctcgg gctatcattgc agcactgggg 12880  
ccagatggta agccctccg tatcgtagtt atctacacga cggggagtcg ggcaactatg 12940  
gatgaacgaa atagacagat cgctgagata ggtgcctcac tgattaagca ttggttaactg 13000  
tcagaccaag ttactcata tatactttag attgatata aacttcattt ttaatttaaa 13060  
aggatctagg tgaagatcct ttttgataat ctcatgacca aaatccctta acgtgagttt 13120  
tcgttccact gagcgtcaga cccgtagaa aagatcaaag gatcttcttg agatccctttt 13180  
tttctgcgcg taatctgctg cttgcaaca aaaaaaccac cgctaccagc ggtggtttgt 13240  
ttgcccgatc aagagctacc aactcttttt ccgaaggtaa ctggcttcag cagagcgag 13300  
ataccaaata ctgtccttct agttagcgg tagttaggcc accacttcaa gaactctgta 13360  
gcaccgccta catacctcgc tctgctaata ctgttaccag tggctgctgc cagtggcgat 13420  
aagtcgtgct ttaccgggtt ggactcaaga cgatagttac cggataaagg gcagcggctg 13480  
ggctgaacgg ggggttcgtg cacacagccc agcttggagc gaacgacctc caccgaactg 13540

37/48

Figure 13T

agatacctac agcgtgagct atgagaaagc gccacgcttc ccgaaggag aaaggcggac 13600  
aggtatccgg taagcggcag ggtcggaaca ggagagcga cgaggagct tccagggga 13660  
aacgcctggt atctttatag tcctgtcggg tttcgccacc tctgacttga gcgtcgattt 13720  
ttgtgatgct cgtcagggg gcggagccta tggaaaaaac ccagcaacgc ggccttttta 13780  
cggttcctgg ccttttgctg cgccgcgtgc ggctgctgga gatggcggac gcgatggata 13840  
tgtttctgcca agggttggtt tgcgcattca cagttctccg caagaattga ttggctccaa 13900  
ttcttggagt ggtgaatccg ttagcgaggt gccgccggct tccattcagg tcgaggtggc 13960  
ccggctccat gcaccgcgac gcaacgcggg gaggcagaca aggtataggc cggcgcctac 14020  
aatccatgcc aaccggttc atgtgctgc cgaggcggca taaatcgccg tgacgatcag 14080  
cggttccagt atcgaagta ggctggtaag agccgcgagc gataccttgaa gctgtcccctg 14140  
atggtcgtca tctacctgcc tggacagcat ggcctgcaac gcgggcatcc cgatgccgcc 14200  
ggaagcgaga agaatacata tggggaaggc catccagcct cgcgtcgcga acgccagcaa 14260  
gacgtagccc agcgcgtcgg ccgccatgcc ctgcttcac cccgtggccc gttgctcgcg 14320  
tttgctggcg gtgtccccg aagaaatata tttgcatgtc tttagttcta tgatgacaca 14380

NruI

38/48

Figure 13U

aaccccgccc agcgtcttgt cattggcgaa ttcgaacacg cagatgcagt cggggcggcg 14440  
cgggtcccagg tcacttcgc atattaaggt gacgcgtgtg gcctcgaaca ccgagcgacc 14500  
ctgcagcgac ccgcttaaca gcgtcaacag cgtgccgcag atcccgggca atgagatatg 14560  
aaaaagcctg aactcacgc gacgtctgtc gagaagtctc tgatcgaaaa gtctcgacagc 14620  
gtctccgacc tgatgcagct ctcgaggggc gaagaatctc gtgctttcag cttcgatgta 14680  
ggagggcgctg gatatgtcct gcgggtaaat agctgcgccg atggtttcta caaagatcgt 14740  
tagtgggac gccactttgc atcgccgcgcg ctccccgatt ccggaagtgc ttgacattgg 14800  
ggaattcagc gagagcctga cctattgcat ctcccgcctg gcacaggggtg tcacgttgca 14860  
agacctgcct gaaaccgaac tgcccgtgtt tctgcagccg gtcgcggagg ccatggatgc 14920  
PvuI gatcgctgcg gccgatctta gccagacgag cgggttcggc ccattcggac cgcaagggaat 14980  
cgggtcaatac actacatggc gtgatttcat atgcgcgatt gctgatcccc atgtgtatca 15040  
ctggcaaaact gtgatggacg acaccgtcag tgcgtccgctc gcgcaggctc tcgatgagct 15100  
gatgctttgg gccgaggact gccccgaagt ccggcacctc gtgcacgcgg atttcggctc 15160  
caacaatgtc ctgacggaca atggccgcat aacagcgggtc attgactgga gcgaggcgat 15220

39/48

Figure 13V

gttcggggat tcccaatacg aggtcgccaa catcttcttc tggaggccgt ggttggcggg 15280  
tatggagcag cagacgcgct acttcgagcg gaggcataccg gagcttgca gatcgccgcg 15340  
gctccggggcg tataatgctcc gcattggtct tgaccaactc tatcagagct tggttgacgg 15400  
caatttcgat gatgcagctt gggcgaggg tcgatgcgac gcaatcgtcc gatccggagc 15460  
cgggactgtc gggcggtacac aaatcgccc cagaagcgcg gccgtctgga ccgatggctg 15520  
tgtagaagta ctgcgccgata gtggaacgg gagatggggg aggtactg aaacacggaa 15580  
ggagacaata ccggaaggaa cccgcgctat gacggcaata aaagacaga ataaacgca 15640  
cgggtgttgg gtcgtttgtt cataaacgcg ggggtcggtc ccagggtgg cactctgtcg 15700  
ataccacc gagaccat tggggccaat acgcccgcgt ttcttcctt tccccacc 15760  
accccccaag ttcgggtgaa ggcccagggc tcgcagccaa cgtcggggcg gcaggccctg 15820  
ccatagccac tggccccgtg ggttagggac ggggtcccc atggggaaatg gtttatgggtt 15880  
cgtgggggtt attatttttg gcgttgctg ggggtctggtc cactgactga ctgagcagac 15940  
agaccatgg ttttggtg gcctgggcat ggaccgcatg tactggcgcg acacgaacac 16000  
cgggcgtctg tggctgcca acacccccga cccccaaaa ccaccgcgcg gatttctggc 16060

40/48

16080

Figure 13W

SalI

gtgccaagct agtcgaccaa





41/48

CTGTTGGGCTCGCGGTTGAGGACAACTCTTCGCGGTCTTTCCAGTACTCTTGGATCGGAAAC  
CCGTCGGCCTCCGAACGGTACTCCGCCACCGAGGGACCTGAGCGAGTCCGCATCGACCGGAT  
CGGAAAACCTCTCGACTGTTGGGGTGAGTACTCCCTCTCAAAAGCGGGCATGACTTCTGCGCT  
AAGATTGTCAGTTTCCAAAAACGAGGAGGATTTGATATTCACCTGGCCCCGCGGTGATGCCTTT  
GAGGGTGGCCGCGTCCATCTGGTCAGAAAAGACAATCTTTTTGTTGTCAAGCTTGAGGTGTGG  
CAGGCTTGAGATCTGGCCATACACTTGAGTGACAATGACATCCACTTTGCCCTTCTCTCCACAG  
GTGTCCACTCCCAGGTCCAACCTGCAGGTCGACTCTAGACCC

FIG. 14A

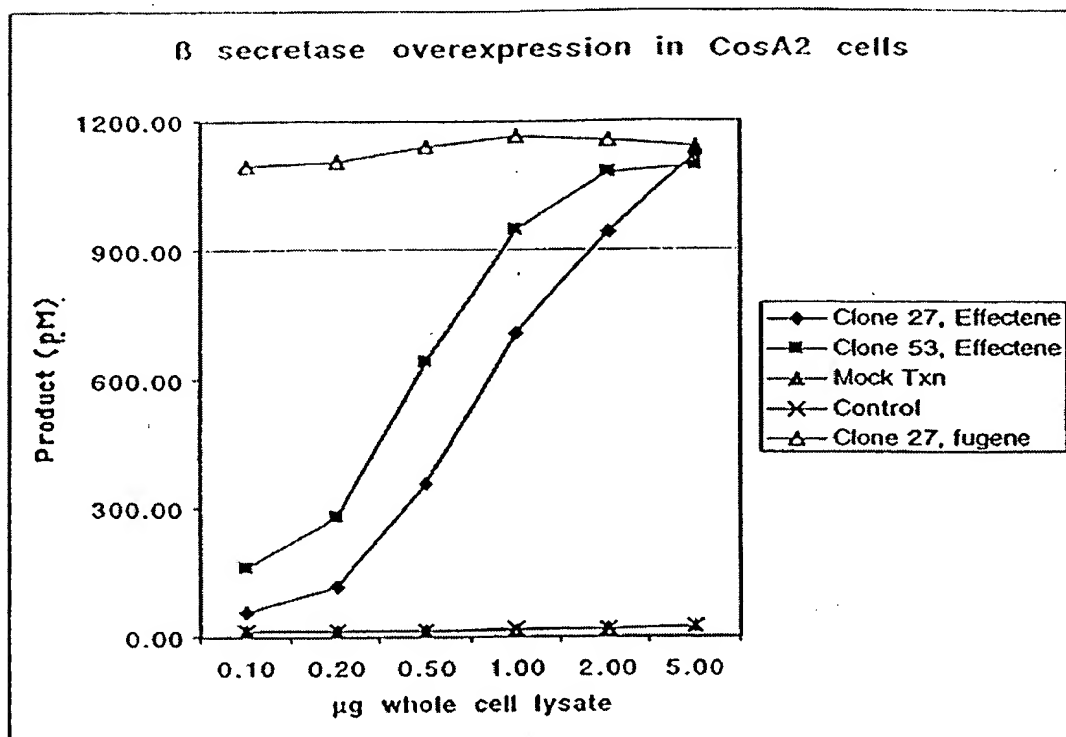


FIG. 14B

42/48

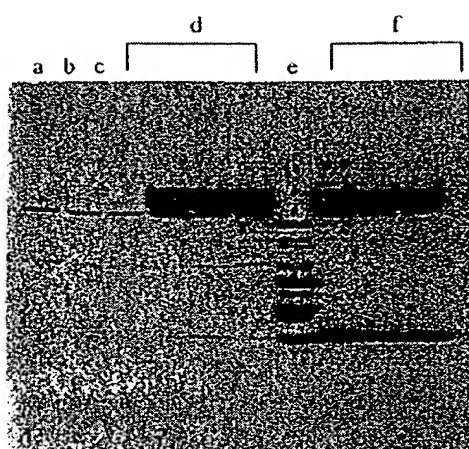


FIG. 15A

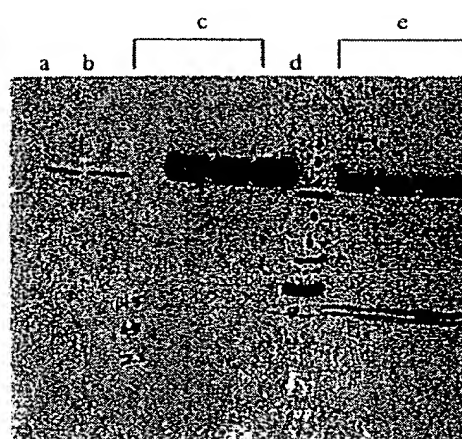


FIG. 15B

43/48

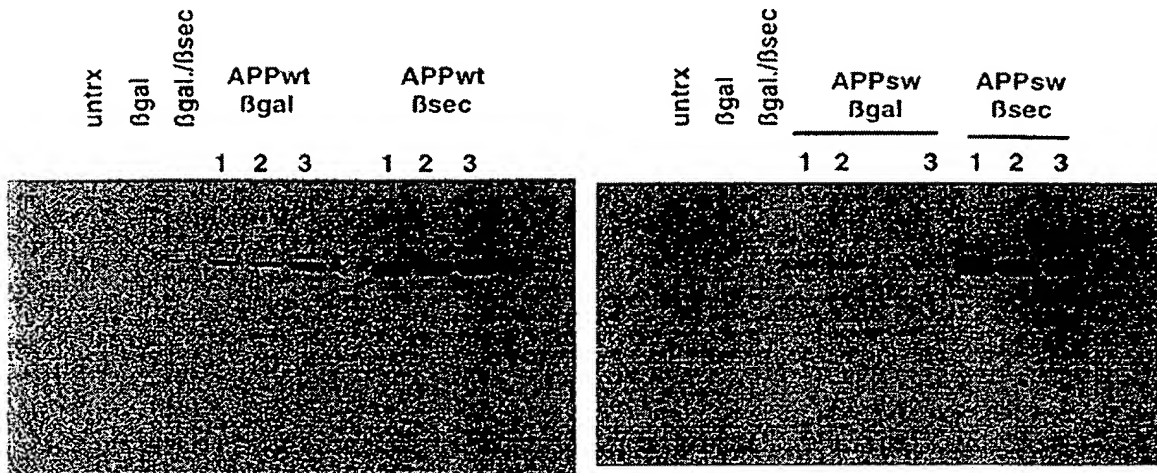


FIG. 16A

FIG. 16B

44/48

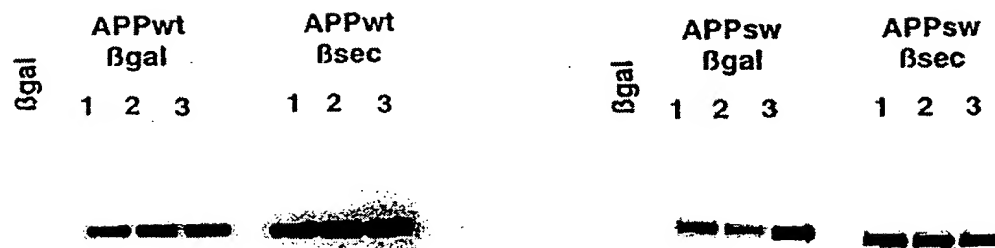


FIG. 17A

FIG. 17B

45/48

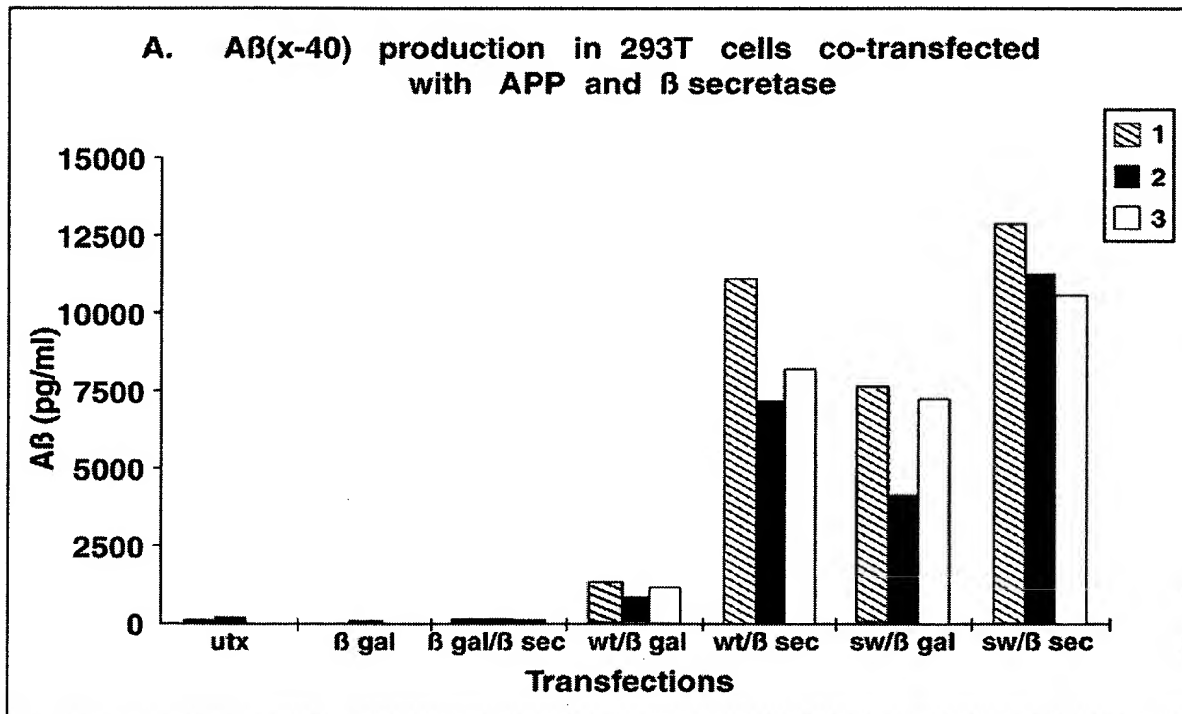


Fig. 18

46/48

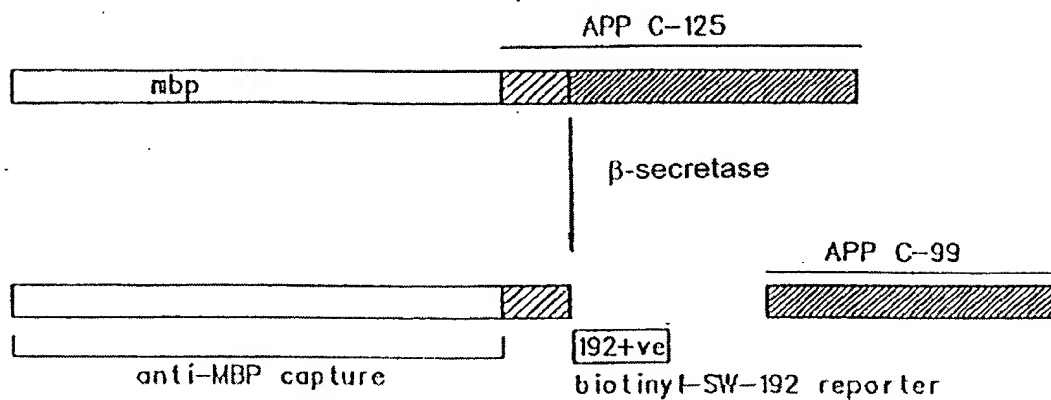


FIG. 19A

Wild-Type Sequence	....Val-Lys-Met-Asp...
Swedish Sequence	....Val-Asn-Leu-Asp...

FIG. 19B

47/48

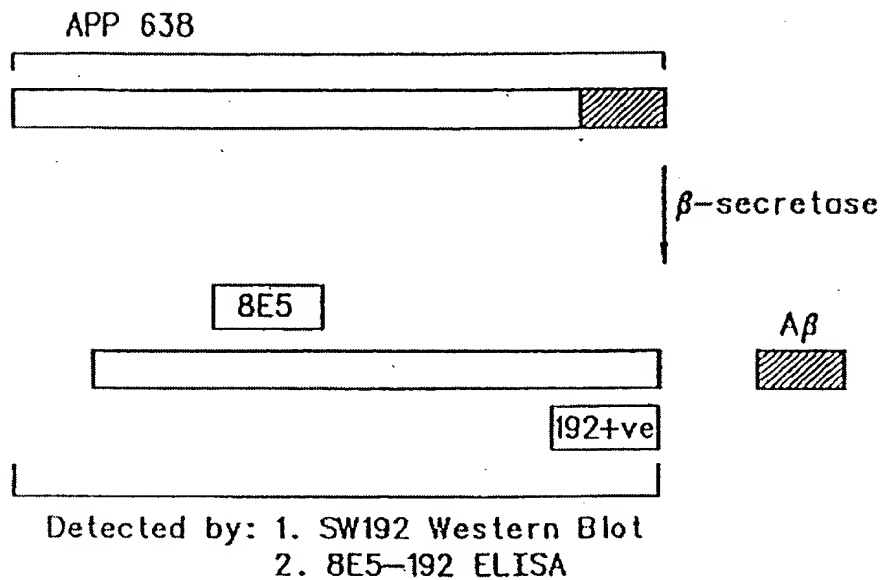


FIG. 20

48/48

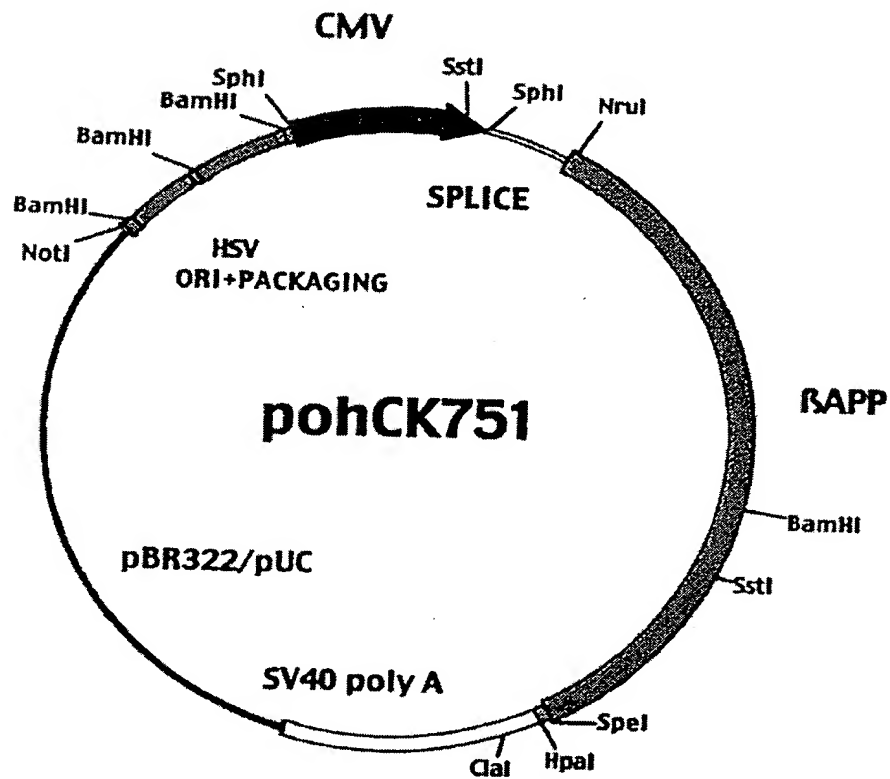


FIG. 21